A Gradient-based Adaptive Algorithm for Minimum Phase - All Pass Decomposition of an FIR system

Mark Flanagan * A. D. Fagan *

Abstract **— An adaptive filtering algorithm which performs minimum phase - all pass (MP-AP) decomposition of an FIR system is proposed. The minimum phase component of the system is modelled as a lattice filter cascaded with a gain stage. Simulation results on randomly generated FIR systems demonstrate convergence of the algorithm to the correct MP-AP decomposition in 92.6% of cases. The algorithm has the additional advantage that it is capable of detecting misconvergence during or after adaptation. These properties combined make the algorithm a strong candidate for use with compound precoding, a pre-equalization method for high-speed modems.**

1 INTRODUCTION

It is well known [1] that any FIR system $G(z)$ of order p may be uniquely decomposed into the cascade of a minimum phase component $W(z)$ (FIR, of order p) and an all pass component. Also, it may be shown [2] that every transversal implementation of a minimum phase FIR filter has an equivalent implementation in the form of a lattice filter followed by a gain stage. The bijective mapping from one filter implementation to the other may be achieved via the Levinson-Durbin and inverse Levinson-Durbin recursions. In this paper we derive and evaluate a gradient adaptive algorithm to achieve this decomposition where the minimum phase component is implemented as a lattice filter followed by a gain stage, and updated accordingly. All stochastic updates are derived by substituting sample correlation for true correlation in the steepest descent updates, as in the ordinary least-mean-square (LMS) algorithm. The lattice structure has the advantage that upon convergence we may easily test whether the filter $W(z)$ given by the algorithm is actually minimum phase. A necessary and sufficient condition for the filter to be minimum phase is that the reflection coefficients of the lattice filter all have magnitude less than unity. Indeed we can check this condition as the algorithm progresses with very little computational effort.

*Department of Electronic and Electrical Engineering, University College Dublin, Ireland, e-mail: mark.flanagan@ee.ucd.ie, tony.fagan@ucd.ie, tel.: +353 1 7161914

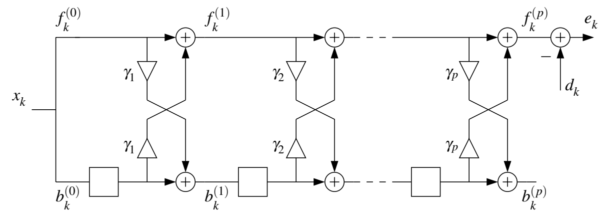

Figure 1: Adaptive lattice filtering for desired response.

2 ADAPTIVE LATTICE FILTERING FOR DESIRED RESPONSE

First we address the general problem of updating the reflection coefficents of a lattice filter in such a way as to obtain minimum mean-squared error (MMSE) between the upper output signal and a desired response. Consider the circuit shown in figure 1. Given the input signal $\{x_k\}$ and the desired response $\{d_k\}$, the problem is to adapt the reflection coefficients $\gamma_j, j = 1, 2 \ldots p$ so as to minimize the mean-squared error (MSE)

$$J(\{\gamma_j\}) = E\left\{e_k^2\right\}$$

where $e_k = f_k^{(p)} - d_k$. In order to derive a stochastic gradient algorithm, we need to estimate $\frac{\partial J}{\partial \gamma_j}$ for each γ_j. It is easy to see that for any $j \in \{1, 2, \ldots p\}$, e_k can be written as $A + B\gamma_j$, where A, B are independent of γ_j. Therefore

$$\begin{aligned}
\frac{\partial J}{\partial \gamma_j} &= \frac{\partial}{\partial \gamma_j} E\left\{(A + B\gamma_j)^2\right\} \\
&= 2E\{AB\} + 2\gamma_j E\left\{B^2\right\} \\
&= 2E\{(A + B\gamma_j)B\} \\
&= 2E\left\{e_k \frac{\partial e_k}{\partial \gamma_j}\right\} \\
&= 2E\left\{e_k \frac{\partial f_k^{(p)}}{\partial \gamma_j}\right\} \\
&= 2E\left\{e_k s_k^{(j)}\right\}
\end{aligned}$$

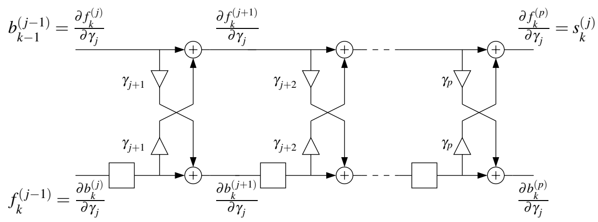

Figure 2: Sublattice for computation of $s_k^{(j)}$, $j \in \{1, 2, \dots p\}$.

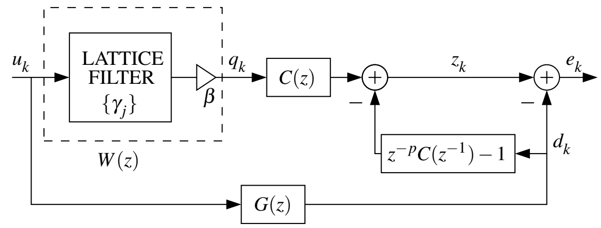

Figure 3: Adaptive algorithm for minimum phase - all pass decomposition of an FIR system $G(z)$.

where we define $s_k^{(j)} = \frac{\partial f_k^{(p)}}{\partial \gamma_j}$ for each $j \in \{1, 2, \dots p\}$. The equations for the lattice filter are

$$\begin{aligned} f_k^{(m)} &= f_k^{(m-1)} + \gamma_m b_{k-1}^{(m-1)} \\ b_k^{(m)} &= \gamma_m f_k^{(m-1)} + b_{k-1}^{(m-1)} \end{aligned} \tag{1}$$

for $m = 1, 2, \dots p$.

Suppose now we pick any $j \in \{1, 2, \dots p\}$. Observe that the sequences $f_k^{(j-1)}$ and $b_{k-1}^{(j-1)}$ are independent of γ_j. Therefore, putting $m = j$ in (1) and differentiating with respect to γ_j,

$$\frac{\partial f_k^{(j)}}{\partial \gamma_j} = b_{k-1}^{(j-1)}$$

and

$$\frac{\partial b_k^{(j)}}{\partial \gamma_j} = f_k^{(j-1)}$$

Also, differentiating (1) with respect to γ_j for any $m \in \{j+1, j+2, \dots p\}$,

$$\begin{aligned} \frac{\partial f_k^{(m)}}{\partial \gamma_j} &= \frac{\partial f_k^{(m-1)}}{\partial \gamma_j} + \gamma_m \frac{\partial b_{k-1}^{(m-1)}}{\partial \gamma_j} \\ \frac{\partial b_k^{(m)}}{\partial \gamma_j} &= \gamma_m \frac{\partial f_k^{(m-1)}}{\partial \gamma_j} + \frac{\partial b_{k-1}^{(m-1)}}{\partial \gamma_j} \end{aligned}$$

This means that each $s_k^{(j)}$ can be computed using a sublattice, as shown in figure 2. Each of these sublattices is attached to the main lattice of figure 1 at the appropriate point. The steepest descent update for the $\{\gamma_j\}$ is then simply

$$\gamma_j^{(k+1)} = \gamma_j^{(k)} - \mu E\left\{e_k s_k^{(j)}\right\}$$

Replacing true correlation with sample correlation gives the adaptive algorithm

$$\gamma_j^{(k+1)} = \gamma_j^{(k)} - \mu e_k s_k^{(j)}$$

Note that the sublattice of figure 2 which computes $s_k^{(j)}$ requires $p-j$ lattice stages. Therefore the entire lattice structure, including sublattices, requires $1 + 2 + \cdots + p = \frac{p(p+1)}{2}$ lattice stages as opposed to the usual p stages.

3 ADAPTIVE ALGORITHM FOR MINIMUM PHASE - ALL PASS DECOMPOSITION

The adaptive circuit for the MP-AP decomposition algorithm is shown in figure 3. The FIR filter $G(z)$ has order p and models the system to be decomposed. The minimum phase component $W(z)$ is implemented as a lattice filter with coefficients $\{\gamma_j\}$ followed by a gain stage β. The IIR section composed of the filters $C(z)$ and $z^{-p}C(z^{-1}) - 1$ represents the all pass component. All filters have order p so as to cater for the possibility of a maximum-phase $G(z)$. The tap-weight $c_p = 1$ and is not updated by the adaptive algorithm. Note that the section comprised of the feedforward and feedback filters with coefficients $\{c_j\}$ is guaranteed to be all pass for any $\{c_j\}$ (and hence at any stage during adaptation) because of the reversal of the coefficients in the feedback section. The MSE is defined as $J(\{c_j\}, \{\gamma_j\}, \beta) = E\{e_k^2\}$. If the convolution of $\mathbf{w} = \begin{pmatrix} w_0 & w_1 & \dots & w_p \end{pmatrix}$ and $\mathbf{c} = \begin{pmatrix} c_0 & c_1 & \dots & c_p \end{pmatrix}$ is denoted $\mathbf{h} = \begin{pmatrix} h_0 & h_1 & \dots & h_{2p} \end{pmatrix}$, then the MSE may be written

$$J = \begin{pmatrix} \mathbf{h} \\ \mathbf{c} \end{pmatrix}^T \mathbf{Q} \begin{pmatrix} \mathbf{h} \\ \mathbf{c} \end{pmatrix}$$

where $\mathbf{Q}$ is a matrix having as entries elements of the autocorrelation sequence $R_u(i) = E\{u_k u_{k-i}\}$ and the cross-correlation sequence $R_{ud}(i) = E\{u_k d_{k-i}\}$. We assume that these two correlation sequences are stationary; hence so is the MSE. From figure 3, the estimation error at time step k may be written as

$$\begin{aligned} e_k &= z_k - d_k \\ &= \sum_{j=0}^{p} c_j q_{k-j} - \sum_{j=0}^{p-1} c_j d_{k-p+j} - d_k \\ &= \sum_{j=0}^{p-1} c_j \left(q_{k-j} - d_{k-p+j}\right) - \left(d_k - q_{k-p}\right) \end{aligned}$$

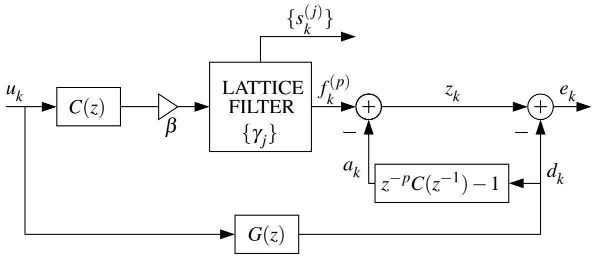

Figure 4: Equivalent circuit to that of figure 3. The lattice filter in this circuit is equipped with sublattices to compute $\left\{s_k^{(j)}\right\}$, $j = 1, 2, \ldots p$.

The sequences $\{q_k\}$ and $\{d_k\}$ are independent of $\{c_j\}$. Therefore the gradient of J with respect to $\{c_j\}$ is given by

$$\left(\frac{\partial J}{\partial c_j}\right)^{(k)} = 2E\left\{e_k\left(q_{k-j} - d_{k-p+j}\right)\right\}$$

Replacing true correlation by sample correlation, we obtain the stochastic gradient update for $\{c_j\}$ as

$$c_j^{(k+1)} = c_j^{(k)} - \mu e_k\left(q_{k-j} - d_{k-p+j}\right)$$

An equivalent circuit to that of figure 3 is shown in figure 4. Since $e_k = f_k^{(p)} - (a_k + d_k)$, the update for the reflection coefficients $\{\gamma_j\}$ is as presented in the discussion of section 2 on adaptive lattice filtering for desired response, *i.e.*

$$\gamma_j^{(k+1)} = \gamma_j^{(k)} - \mu e_k s_k^{(j)} \qquad j = 1, 2, \ldots p$$

where $s_k^{(j)}$ is computed by the sublattices. To derive the update for the gain parameter β, note that if the order of the lattice filter and gain stage were switched in figure 4, the input to the gain stage would be $f_k^{(p)}/\beta$. Therefore the update for the gain parameter β may be obtained as simply the stochastic gradient update for a one-tap FIR filter, *i.e.*

$$\beta^{(k+1)} = \beta^{(k)} - \mu e_k\left(\frac{f_k^{(p)}}{\beta}\right)$$

Finally note that if the circuit of figure 4 is implemented, the entire equivalent circuit of figure 3 need not be. The lattice filter (without sublattices) and gain stage, followed by a p-element delay line is sufficient.

4 ALGORITHM PERFORMANCE

The adaptive algorithm was tested on a set of 5000 filters $G(z)$ of order $p = 9$, the coefficients of $G(z)$ in each case being taken from a uniform probability distribution on the interval $[-1, 1]$. Since the error surface is multimodal the choice of initial filter coefficients is important (there are in general 2^p valid MMSE solutions, only one of which gives a minimum phase $W(z)$). The reflection coefficients were initialized to $\gamma_j^{(0)} = 0$ for $j = 1, 2, \ldots p$; this filter is "minimum phase optimum" in the sense that its zeros are at maximum distance from the unit circle while maintaining minimum phase. The gain parameter was initialized to $\beta^{(0)} = 10$. Small values of $\beta^{(0)}$ (~ 1) were found to result in more probable convergence to a non minimum phase $W(z)$, while large values of $\beta^{(0)}$ (~ 100) were found to result in numerical instability. $C(z)$ was initialized as a comb filter, *i.e.* $c_0^{(0)} = c_p^{(0)} = 1$ and $c_j^{(0)} = 0$ for $j \notin \{0, p\}$, so that at the beginning the all pass section has cancelling pole-zero pairs evenly spaced on the unit circle. The excitation $\{u_k\}$ was chosen to be a white sequence (with elements equiprobable in $\{-1, +1\}$) in order to identify the system $G(z)$ at all frequencies. The step size was chosen to be $\mu = 10^{-3}$, small enough so that of the 5000 simulation runs, each of 20,000 iterations, none diverged.

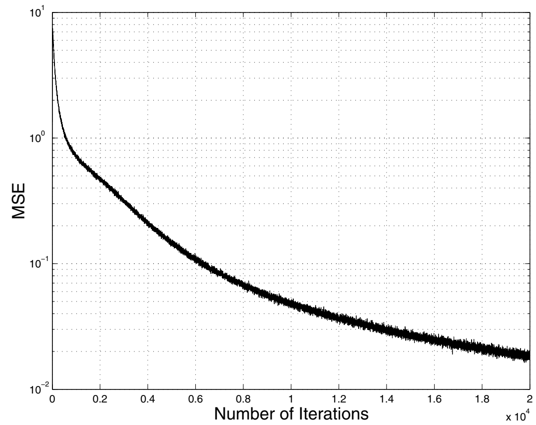

Figure 5: MSE convergence, ensemble averaged over 5000 simulations of the adaptive algorithm.

As the filter $W(z)$ converges to the minimum phase component of $G(z)$, the zeros of the all pass filter move outwards from the unit circle and the poles of the all pass filter (which are the reciprocals of its zeros) move inside the unit circle to provide the correct compensation. It was found that in 4628 of the 5000 cases, the filter given by the algorithm was minimum phase (92.6% success rate), and the average final MSE was 1.9×10^{-2}. Figure 5 shows the evolution of the ensemble averaged MSE with the number of iterations. Also, by way of illustration, z-plane results are shown in figures 6, 7 and

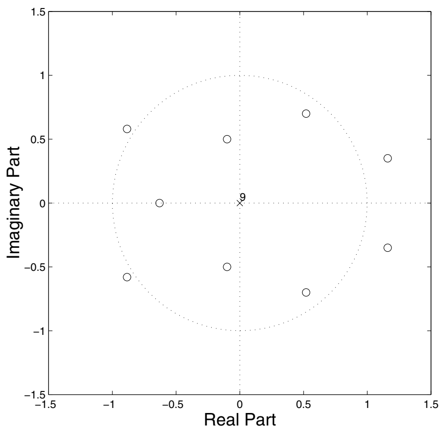

Figure 6: Pole-zero plot for the original filter $G(z)$.

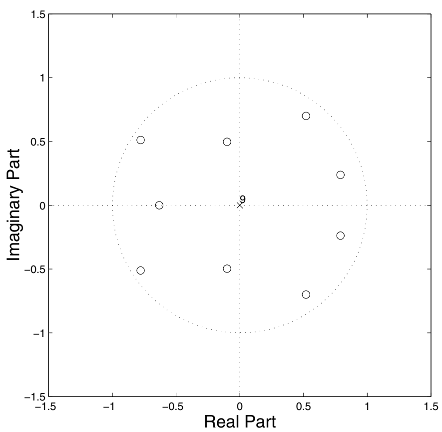

Figure 7: Pole-zero plot for the minimum phase component of $G(z)$ given by the adaptive algorithm.

8 for one of these filters $G(z)$. Figure 6 shows the pole-zero diagram of the original filter to be decomposed. Figure 7 shows the pole-zero diagram of the minimum phase component given by the adaptive algorithm, and figure 8 shows the pole-zero diagram for the all pass component given by the adaptive algorithm.

5 APPLICATION: COMPOUND PRECODING FOR V.92

Compound precoding is a power efficient method for combining decision-feedback equalization (DFE) with trellis coding in a modem transmitter, and is an option in the V.92 standard [3, 4]. To realize the gain of compound precoding however, the feedforward filter of the compound precoder must be decomposed into its minimum phase and

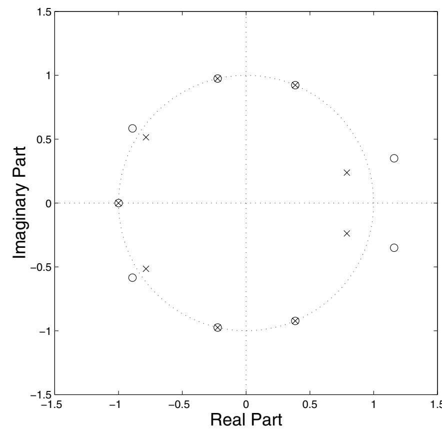

Figure 8: Pole-zero plot for the all pass component of $G(z)$ given by the adaptive algorithm.

all pass components [3]. The adaptive algorithm proposed here is a strong candidate for performing this decomposition, for two reasons: firstly, it has a high probability of convergence to the correct MP-AP decomposition, and secondly, misconvergence of the algorithm may be detected at any stage during adaptation and another method of precoding selected.

6 CONCLUSION

We have proposed and evaluated an adaptive filtering algorithm which performs MP-AP decomposition of an FIR system. The minimum phase component of the system is modelled as a lattice filter cascaded with a gain stage. With this adaptive filter structure, algorithm misconvergence is of low probability and is immediately detectable. These properties make the algorithm suitable for use with compound precoding.

References

[1] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ.: Prentice Hall, 1989.

[2] S. Haykin, *Adaptive Filter Theory*. Information and System Science Series, Prentice Hall, 1996.

[3] M. McLaughlin, “Compound Precoding: Pre-equalization for channels with combined feed-forward and feedback characteristics”, *Irish Signals and Systems Conference*, pp. 536–543, 2000.

[4] ITU-T Recommendation V.92, “Enhancements to recommendation V.90”, 2000.